

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Stiles, Michael E.
Verderas, John C.
Van Belkum, Marius J.
Worobo, Randy W.
Worobo, Rodney J.
McCormick, John K.
McMullen, Lynn M.
Leisner, Jorgen J.
Poon, Alison
Franz, Charles MAP

(ii) TITLE OF INVENTION: NOVEL BACTERIOCINS, TRANSPORT AND VECTOR
SYSTEM AND METHOD OF USE THEREOF

(iii) NUMBER OF SEQUENCES: 52

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Richardson
(B) STREET: 2200 Sand Hill Road
(C) CITY: Menlo Park
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94025

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Green, Grant D.
(B) REGISTRATION NUMBER: 31,259
(C) REFERENCE/DOCKET NUMBER: 07254/051001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-233-2126
(B) TELEFAX: 415-854-0875

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 184 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: RBS
(B) LOCATION: 24..28

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 38..184

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 38..125

(ix) FEATURE:

(A) NAME/KEY: primer_bind
(B) LOCATION: 114..144

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION: 114..119
(D) OTHER INFORMATION: /label= Restrict_site

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATCTTGATAT CACAACTAA TTTGGAGGTT GGTATAT ATG AAA AAA CAA ATT TTA 55
Met Lys Lys Gln Ile Leu
1 5

AAA GGG TTG GTT ATA GTT GTT TGT TTA TCT GGG GCA ACA TTT TTC TCA 103
Lys Gly Leu Val Ile Val Val Cys Leu Ser Gly Ala Thr Phe Phe Ser
10 15 20

ACA CCA CAA CAA GCT TCT GCT GTA AAT TAT GGT AAT GGT GTT TCT TGC 151
Thr Pro Gln Gln Ala Ser Ala Val Asn Tyr Gly Asn Gly Val Ser Cys
25 30 35

AGT AAA ACA AAA TGT TCA GTT AAC TGG GGA CAA 184
Ser Lys Thr Lys Cys Ser Val Asn Trp Gly Gln
40 45

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10 Met Lys Lys Gln Ile Leu Lys Gly Leu Val Ile Val Val Cys Leu Ser
1 5 10 15
Gly Ala Thr Phe Phe Ser Thr Pro Gln Gln Ala Ser Ala Val Asn Tyr
20 25 30
15 Gly Asn Gly Val Ser Cys Ser Lys Thr Lys Cys Ser Val Asn Trp Gly
35 40 45

Gln

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: Leucocin A genes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

40 AATTTTGGCC CATGCTGCAT GATATTTTGA CTACCAAAAA ATATGCGTGT TGCCTACTTC 60
AATGTTGATA ATTTTTTTAA AGATAATTCC TCTGACAAAG CTAGTTATAT TAATTTCTTT 120
CAAGAGTTAA ATATTTCTCA ACTGCCTTCT TTAATTTTTA CTAATGGAAA CATGGACTAT 180
45 AAACGATTAT CAATTTATAC AATTAAAACA CCAATAAATG CATGGATTAC TGCTATTAAT 240
GACGAATTAA TTTCAAACA TTCCAAGCAA TCATCAACAA ATTAAAAATG GTTAAGGTCA 300
AAATGTTTCA AAAAAGAATA AATTATATCG CACAAGTAGA TGAACGTGAT TGTGGTGTG 360
50 CTGCACTCGC TATGGTTTAA ACTCATTACA AAACACGCCT GTCCTTAGCC AACTACGGG 420
ACCTGGCCAA AACTGACATG GAAGGAACGA CTGCTTTAGG CATTGTTAAA GCTGCGAATG 480

	CGCTAGACTT TGAAACCATG CCGATCCAGG CTGATTTGAG TTTATTCGAT AAAAAGGATT	540
	TACCCTATCC TTTTATCGCC CATGTCATTA AAGACGGTAA ATACCCGCAT TATTATGTAG	600
5	TTTATGGGAT CAAAGGTGAT CAGCTATTAA TCGCTGATCC AGATAATACC GTTGGTAAAA	660
	ATAAAATGAC AAAAGCGCAT TTTAATGAGG AGTGGACCGG TGTGTCCATT TTTATTGCGC	720
	CCAATCCAAC CTACAAGCCA ACAAAGGATA AAAAGCGTTC CTTGACTTCT TTTATTCCAG	780
10	TGATTACGCG TCAAAAATTA TTAGTTATCA ATATTGTCAT TGCTGCCTTG TTGGTTACCC	840
	TAGTGAGTAT TTTAGGATCA TATTATTTGC AAGGTATCAT TGATACCTAT ATCCCCGATA	900
15	ACATGAAAAA CACCCTAGGG ATTGTGTCAC TAGGGCTTAT TTTTGCGTAT GTTATCCAAC	960
	AAGTGTCTC TTATGCCAGA GATTATTTAT TAATTGTCAT GGGGCAACGC CTCTCAATTG	1020
	ATATTATTTT GTCTTATATC AAACACATTT TTGAACTGCC AATGTCTTTT TTCGCGACGC	1080
20	GTCGTACCGG TGAAATTGTG AGCCGTTTTA CGGACGCTAA TGCCATTATT GAAGCCCTGG	1140
	CAAGCACGAT GTTATCTGTA TTTTITAGACT TAGGAATTTT GGTCATTGTT GGCACAGTGC	1200
25	TAGTGGTTCA AAATTCAACC TTGTTTCTGA TTTCTCTGAT TGCCATTCCG GCTTATGCCC	1260
	TAGTGGTCTG GCTCTTTATG CGTCCTTTTT CAAAGATGAA TAATGACCAA ATGCAAGCAG	1320
	GTTGATGTT AAGTTCTTCC ATTATTGAAG ATATTAATGG CGTTGAGACG ATTAAAGCGC	1380
30	TGAATAGTGA AGAAACCGCC TATCATAAAA TTGATCATGA ATTTGTCACT TATTTAGAAA	1440
	AATCATTTGT TTACGCTAAA ACAGAAGCCA CTCAAATGC GATTAAAAGC CTCTTACAGC	1500
35	TCTCTTTAAA TGTCGTGATC TTATGGGTTG GCGCACAACT GGTCATGACC AATAAAATTA	1560
	GTGTTGGTCA ACTGATCACT TACAATGCTT TATTAGGATT TTTTACAGAT CCCTTGCAAA	1620
	ATATTATTAA TTTACAAACT AAGCTCCAAC AGGCCTCAGT CGCTAATAAT CGTTTGAACG	1680
40	AAGTTTATTT GGTTGATTCA GAATTTAAAG CTAGTCATCA AATGACAGAA AGCATTATGC	1740
	CCAATAGCTC ATTAGTAGCC GATCATATCA CCTATAAATA CGGTTTTGGT GCGCCAGCAA	1800
45	TTGATGATGT TTCACTAACG ATTACAGCCG GTGAAAAAAT CGCTTTGGTT GGGATTAGTG	1860
	GATCAGGTAA ATCAACTTTA GTTAAATTGC TGGTTAATTT CTTTCAACCA GAGTCAGGGA	1920
	CAATTTCACT AGGACAAACA CCACTCGCCA ATCTTGATAA ACATGAGCTA AGAGCACACA	1980
50	TTAATTATTT ACCACAAGAA CCCTTTATAT TTTCCGGTTC AATTATGGAC AACCTGTTAT	2040
	TGGGGGCTAA GCCAGGGACA ACCCAAGAAG ATATTATCAG GGCGGTAGAA ATTGCTGAAA	2100

	TTAAAGATGA TATTGAAAAA ATGTCGCAAG GATTTGGCAC TGAAGTCGCA GAAAGTGGCA	2160
	ATATTTTCGGG TGGTCAAAAA CAACGCATTG CTTTAGCTAG AGCCATTTTA GTCGATTCTC	2220
5	CGGTGCTGAT TTTAGATGAG TCAACCAGTA ATCTTGATGT TTTAACAGAA AAAAAGATTA	2280
	TTGATAATCT CATGCAGTTA ACCGAAAAAA CCATTATCTT TGTAGCGCAC CGCTTAACCA	2340
	TTTCACAGCG AGTAGATCGT ATTCTAACCA TGCAAAACGG CAAAATTATC GAAGATGGCA	2400
10	CGCATAATAC TCTGCTTAAT GCCGGTGGTT TCTACGCGTC ATTGTTTAAT CATTAAAGGAG	2460
	ACCTGATGTT TGATCCAAAA TACTTAGAAA GTGGCGAATT TTATCAACGT CGTTACCGCA	2520
15	ATTTTCCAAC TCTGATTATT GTGCCTATTT TTTTGTTAGT CGTGTTTATC ATTCTATTTA	2580
	GCCTATTTGC TAAGCGTGAA ATTGTTGTCA AAGCAAGTGG CGAAATTATT CCAGCCAAAG	2640
	TGCTATCAGA TATCCAATCA ACCAGTAACA ATGCCATCGA TAGTAACCAA TTAAGTAAAA	2700
20	ATAAAGTGGT TAAAAAGGC GATACCTTAG TGACCTTTAC CAGTGGTAAT GAAAAAATAT	2760
	CGTCTCAATT ACTGACGCAA CAACTTAATA ATCTTAACGA CCGTCTAAAA AGTCTTGATA	2820
25	CCTATAAGCA GAGTATTGTT AACGGACGTA GCGAATTTGG TGGCACAGAT CAATTTGGTT	2880
	ATGATAGTCT ATTCAACGGC TATATGGCGC AAGTTGATAC GTTGACGAGT GAATTTAATC	2940
	AACAAAGTAG TGATAAACAA ACAGCTGATC AACAAGCTAA TCATCAAATT GACGTTTTAA	3000
30	AACAAGGTCA ATCTAAAAAC AATCAACAAT TAGCTAATTA TCAAGCTATT CTAACCAGTA	3060
	TTAATAGCAA CACTAAACCG ACTAATAATC CCTATCAAGC CATTTATGAT AATTATTCAG	3120
35	CCCAGTTAAA ATCAGCACAA ACAACTGATG ATAAAGATCA AGTCAAGCAA ACTGCCTTAA	3180
	GTAATGTACA ACAACAAATT GATCAATTAC AAACAACGAG TAGTTCGTAT GATAGTCAAA	3240
	TTGCTGGTAT TACAAAGAGT GGTCCTTTAT CTCAAAGCAG TACCTTAGAT AAAATCGCTG	3300
40	ACTTGAAGCA ACAACAATA GCGAGTGCTC AAAAAGAAAT CAATGATCAG CAACAATCCT	3360
	TAGATGAGTT AAAAGCCAAG CAATCCTCTG CTAATGAGGA TTATCAAGAT ACGGTTATTA	3420
45	AAGCACCAGA AGATGGCATT TTACATTTAG CCACTGACAA AACTAAAATC AAGTATTTCC	3480
	CTAAAGGCAC AACCATTGCG CAAATTTATC CTAACTGAC GCAAAAAACA GCTTTGAATG	3540
	TTGAGTACTA TGTGCCTGCC AGTAATATTA TCGGCTTAAA GCAAAGACAA GCCATCCGTT	3600
50	TTGTAGCAAA TCAAAATGTC ACGAAACCGC TCACCTTAAA CGGAACAATC AAAAGCATTA	3660
	GTTCTGCACC AATAGCCAGT AAAGAGGGAT CCTTTTATAA ATTAGTCGCG ACGATTCAGG	3720

CTAGCAAAAT AGACCGTGAA CAGATTAAAT ATGGTCTTAA TGGTCGAATC ACAACCATAA 3780
AAGGGACTAA AACATGGTTT AATTATTATA AAGACATTGT TTTAGGTGAG AATAATTAGC 3840
5 TAGGAAGATA AACACAATTT TTAAACGTGT TTATCTTTTT TAGTCTCAAT GAAATTGTCTG 3900
CCGAAGGTTT TTCTAGCCAA GTGGCAGGAC ACAGAAAAAT GATAGTTGCT ACTGAAGGGA 3960
AGTTCAACTG CCACCAAAAA TAGTAACCGC GCGACAGCCA ACCGCCACCA CAACAGTTAT 4020
10 GCTCGCCCGT GGTTATTATT ATCATTAACA CTCTTACGTC TTTCTATGAT ACTTTTGAGC 4080
CACATTCTTA TAATGCTGCA ATCGACCTTT TAGAAAATTG ATCTCATCAG AAATTTCTTT 4140
15 TAAGTGGTTA TCATCAGCAT GTTACTAGC AATATTTAAT TCTTTAATCC TACGTTTAAT 4200
CAACTTAGTA GTTTTAGTAT CTTTCATGTA TTGATTATCT CAAAAAACA CCCAACAAGG 4260
GCAATCAGTT TGATTTGAGC AGAGGAAGCC 4290

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 717 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: lcaC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Phe	Gln	Lys	Arg	Ile	Asn	Tyr	Ile	Ala	Gln	Val	Asp	Glu	Arg	Asp
1				5					10					15	
Cys	Gly	Val	Ala	Ala	Leu	Ala	Met	Val	Leu	Thr	His	Tyr	Lys	Thr	Arg
			20					25					30		
Leu	Ser	Leu	Ala	Lys	Leu	Arg	Asp	Leu	Ala	Lys	Thr	Asp	Met	Glu	Gly
			35				40					45			
Thr	Thr	Ala	Leu	Gly	Ile	Val	Lys	Ala	Ala	Asn	Ala	Leu	Asp	Phe	Glu
			50			55				60					
Thr	Met	Pro	Ile	Gln	Ala	Asp	Leu	Ser	Leu	Phe	Asp	Lys	Lys	Asp	Leu

65	70	75	80
Pro Tyr Pro Phe	Ile Ala His Val	Ile Lys Asp Gly Lys Tyr	Pro His
	85	90	95
Tyr Tyr Val Val	Tyr Gly Ile Lys	Gly Asp Gln Leu Leu	Ile Ala Asp
	100	105	110
Pro Asp Asn Thr	Val Gly Lys Asn	Lys Met Thr Lys	Ala His Phe Asn
	115	120	125
Glu Glu Trp Thr	Gly Val Ser Ile	Phe Ile Ala Pro	Asn Pro Thr Tyr
	130	135	140
Lys Pro Thr Lys	Asp Lys Lys Arg	Ser Leu Thr Ser	Phe Ile Pro Val
	145	150	155
Ile Thr Arg Gln	Lys Leu Leu Val	Ile Asn Ile Val	Ile Ala Ala Leu
	165	170	175
Leu Val Thr Leu	Val Ser Ile Leu	Gly Ser Tyr Tyr	Leu Gln Gly Ile
	180	185	190
Ile Asp Thr Tyr	Ile Pro Asp Asn	Met Lys Asn Thr	Leu Gly Ile Val
	195	200	205
Ser Leu Gly Leu	Ile Phe Ala Tyr	Val Ile Gln Gln	Leu Leu Ser Tyr
	210	215	220
Ala Arg Asp Tyr	Leu Leu Ile Val	Met Gly Gln Arg	Leu Ser Ile Asp
	225	230	235
Ile Ile Leu Ser	Tyr Ile Lys His	Ile Phe Glu Leu	Pro Met Ser Phe
	245	250	255
Phe Ala Thr Arg	Arg Thr Gly Glu	Ile Val Ser Arg	Phe Thr Asp Ala
	260	265	270
Asn Ala Ile Ile	Glu Ala Leu Ala	Ser Thr Met Leu	Ser Val Phe Leu
	275	280	285
Asp Leu Gly Ile	Leu Val Ile Val	Gly Thr Val Leu	Val Val Gln Asn
	290	295	300
Ser Thr Leu Phe	Leu Ile Ser Leu	Ile Ala Ile Pro	Ala Tyr Ala Leu
	305	310	315
Val Val Trp Leu	Phe Met Arg Pro	Phe Ser Lys Met	Asn Asn Asp Gln
	325	330	335
Met Gln Ala Gly	Ser Met Leu Ser	Ser Ser Ile Ile	Glu Asp Ile Asn
	340	345	350
Gly Val Glu Thr	Ile Lys Ala Leu	Asn Ser Glu Glu	Thr Ala Tyr His

355 360 365

Lys Ile Asp His Glu Phe Val Thr Tyr Leu Glu Lys Ser Phe Val Tyr
370 375 380

Ala Lys Thr Glu Ala Thr Gln Asn Ala Ile Lys Ser Leu Leu Gln Leu
385 390 395 400

Ser Leu Asn Val Val Ile Leu Trp Val Gly Ala Gln Leu Val Met Thr
405 410 415

Asn Lys Ile Ser Val Gly Gln Leu Ile Thr Tyr Asn Ala Leu Leu Gly
420 425 430

Phe Phe Thr Asp Pro Leu Gln Asn Ile Ile Asn Leu Gln Thr Lys Leu
435 440 445

Gln Gln Ala Ser Val Ala Asn Asn Arg Leu Asn Glu Val Tyr Leu Val
450 455 460

Asp Ser Glu Phe Lys Ala Ser His Gln Met Thr Glu Ser Ile Met Pro
465 470 475 480

Asn Ser Ser Leu Val Ala Asp His Ile Thr Tyr Lys Tyr Gly Phe Gly
485 490 495

Ala Pro Ala Ile Asp Asp Val Ser Leu Thr Ile Thr Ala Gly Glu Lys
500 505 510

Ile Ala Leu Val Gly Ile Ser Gly Ser Gly Lys Ser Thr Leu Val Lys
515 520 525

Leu Leu Val Asn Phe Phe Gln Pro Glu Ser Gly Thr Ile Ser Leu Gly
530 535 540

Gln Thr Pro Leu Ala Asn Leu Asp Lys His Glu Leu Arg Ala His Ile
545 550 555 560

Asn Tyr Leu Pro Gln Glu Pro Phe Ile Phe Ser Gly Ser Ile Met Asp
565 570 575

Asn Leu Leu Leu Gly Ala Lys Pro Gly Thr Thr Gln Glu Asp Ile Ile
580 585 590

Arg Ala Val Glu Ile Ala Glu Ile Lys Asp Asp Ile Glu Lys Met Ser
595 600 605

Gln Gly Phe Gly Thr Glu Leu Ala Glu Ser Gly Asn Ile Ser Gly Gly
610 615 620

Gln Lys Gln Arg Ile Ala Leu Ala Arg Ala Ile Leu Val Asp Ser Pro
625 630 635 640

Val Leu Ile Leu Asp Glu Ser Thr Ser Asn Leu Asp Val Leu Thr Glu

645 650 655
Lys Lys Ile Ile Asp Asn Leu Met Gln Leu Thr Glu Lys Thr Ile Ile
660 665 670
Phe Val Ala His Arg Leu Thr Ile Ser Gln Arg Val Asp Arg Ile Leu
675 680 685
Thr Met Gln Asn Gly Lys Ile Ile Glu Asp Gly Thr His Asn Thr Leu
690 695 700
Leu Asn Ala Gly Gly Phe Tyr Ala Ser Leu Phe Asn His
705 710 715

15 (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 457 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: lcaD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Phe Asp Pro Lys Tyr Leu Glu Ser Gly Glu Phe Tyr Gln Arg Arg
1 5 10 15
Tyr Arg Asn Phe Pro Thr Leu Ile Ile Val Pro Ile Phe Leu Leu Val
20 25 30
Val Phe Ile Ile Leu Phe Ser Leu Phe Ala Lys Arg Glu Ile Val Val
35 40 45
Lys Ala Ser Gly Glu Ile Ile Pro Ala Lys Val Leu Ser Asp Ile Gln
50 55 60
Ser Thr Ser Asn Asn Ala Ile Asp Ser Asn Gln Leu Thr Glu Asn Lys
65 70 75 80
Val Val Lys Lys Gly Asp Thr Leu Val Thr Phe Thr Ser Gly Asn Glu
85 90 95
Lys Ile Ser Ser Gln Leu Leu Thr Gln Gln Leu Asn Asn Leu Asn Asp

100 105 110

Arg Leu Lys Ser Leu Asp Thr Tyr Lys Gln Ser Ile Val Asn Gly Arg
115 120 125

5 Ser Glu Phe Gly Gly Thr Asp Gln Phe Gly Tyr Asp Ser Leu Phe Asn
130 135 140

10 Gly Tyr Met Ala Gln Val Asp Thr Leu Thr Ser Glu Phe Asn Gln Gln
145 150 155 160

Ser Ser Asp Lys Gln Thr Ala Asp Gln Gln Ala Asn His Gln Ile Asp
165 170 175

15 Val Leu Lys Gln Gly Gln Ser Lys Asn Asn Gln Gln Leu Ala Asn Tyr
180 185 190

Gln Ala Ile Leu Thr Ser Ile Asn Ser Asn Thr Lys Pro Thr Asn Asn
195 200 205

20 Pro Tyr Gln Ala Ile Tyr Asp Asn Tyr Ser Ala Gln Leu Lys Ser Ala
210 215 220

Gln Thr Thr Asp Asp Lys Asp Gln Val Lys Gln Thr Ala Leu Ser Asn
225 230 235 240

25 Val Gln Gln Gln Ile Asp Gln Leu Gln Thr Thr Ser Ser Ser Tyr Asp
245 250 255

30 Ser Gln Ile Ala Gly Ile Thr Lys Ser Gly Pro Leu Ser Gln Ser Ser
260 265 270

Thr Leu Asp Lys Ile Ala Asp Leu Lys Gln Gln Gln Leu Ala Ser Ala
275 280 285

35 Gln Lys Glu Ile Asn Asp Gln Gln Gln Ser Leu Asp Glu Leu Lys Ala
290 295 300

Lys Gln Ser Ser Ala Asn Glu Asp Tyr Gln Asp Thr Val Ile Lys Ala
305 310 315 320

40 Pro Glu Asp Gly Ile Leu His Leu Ala Thr Asp Lys Thr Lys Ile Lys
325 330 335

45 Tyr Phe Pro Lys Gly Thr Thr Ile Ala Gln Ile Tyr Pro Lys Leu Thr
340 345 350

Gln Lys Thr Ala Leu Asn Val Glu Tyr Tyr Val Pro Ala Ser Asn Ile
355 360 365

50 Ile Gly Leu Lys Gln Arg Gln Ala Ile Arg Phe Val Ala Asn Gln Asn
370 375 380

Val Thr Lys Pro Leu Thr Leu Asn Gly Thr Ile Lys Ser Ile Ser Ser

[illegible]

15 (2) INFORMATION FOR SEQ ID NO:6:

- ```

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:
 (B) CLONE: Divergicin N-terminal

(ix) FEATURE:
 (A) NAME/KEY: Cleavage-site
 (B) LOCATION: 29..30

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

[illegible]

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 amino acids

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:  
(B) CLONE: Divergicin N-terminal, cleaved

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Lys Lys Gln Ile Leu Lys Gly Leu Val Ile Val Val Cys Leu Ser  
1 5 10 15

Gly Ala Thr Phe Phe Ser Thr Pro Gln Gln Ala Ser Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:  
(B) CLONE: leucocin A N-terminal

(ix) FEATURE:  
(A) NAME/KEY: Cleavage-site  
(B) LOCATION: 24..25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Met Asn Met Lys Pro Thr Glu Ser Tyr Glu Gln Leu Asp Asn Ser  
1 5 10 15

Ala Leu Glu Gln Val Val Gly Gly Lys Tyr Tyr Gly Asn Gly  
20 25 30

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:

(B) CLONE: leucocin A, cleaved

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Met Asn Met Lys Pro Thr Glu Ser Tyr Glu Gln Leu Asp Asn Ser  
1 5 10 15

Ala Leu Glu Gln Val Val Gly Gly  
20

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:

(B) CLONE: lactococcin A N-terminal

(ix) FEATURE:

(A) NAME/KEY: Cleavage-site

(B) LOCATION: 21..22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Lys Asn Gln Leu Asn Phe Asn Ile Val Ser Asp Glu Glu Leu Ser  
1 5 10 15

Glu Ala Asn Gly Gly Lys Leu Thr Phe Ile Gln  
20 25

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:

(B) CLONE: lactococcin A, cleaved

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Lys Asn Gln Leu Asn Phe Asn Ile Val Ser Asp Glu Glu Leu Ser  
1 5 10 15

Glu Ala Asn Gly Gly  
20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:  
(B) CLONE: colicin V, N-terminal

(ix) FEATURE:  
(A) NAME/KEY: Cleavage-site  
(B) LOCATION: 15..16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Arg Thr Leu Thr Leu Asn Glu Leu Asp Ser Val Ser Gly Gly Ala  
1 5 10 15

Ser Gly Arg Asp Ile  
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vii) IMMEDIATE SOURCE:  
(B) CLONE: colicin V, cleaved

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Arg Thr Leu Thr Leu Asn Glu Leu Asp Ser Val Ser Gly Gly  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3475 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: pCD3.4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| GATATCTTGG | TATTACAAAC | TAATTGGAGG | TTGGTATATA | TGAAAAAACA | AATTTTAAAA | 60   |
| GGGTTGGTTA | TAGTTGTTTG | TTTATCTGGG | GCAACATTTT | TCTCAACACC | ACAACAAGCT | 120  |
| TCTGCTGCTG | CACCGAAAAT | TACTCAAAAA | CAAAAAAATT | GTGTTAATGG | ACAATTAGGT | 180  |
| GGAATGCTTG | CTGGAGCTTT | GGGTGGACCT | GGCGGAGTTG | TGTTAGGTGG | TATAGGTGGT | 240  |
| GCAATAGCAG | GAGGTTGTTT | TAATTAAATG | AAAATAAAAT | GGTACTGGGA | ATCTCTGATT | 300  |
| GAAACCTTAA | TATTTATAAT | TGTTCTTCTT | GTATTTTTTT | ATAGAAGTTC | TGGTTTTTCT | 360  |
| TTAAAAAATT | TAGTTTTAGG | AAGTTTATTT | TATTTGATAG | CAATTGGTCT | TTTTAATTAT | 420  |
| AAAAAGATAA | ACAAATAGGC | ACTATTTTTA | AATTTACAAC | TTTTGCATTT | TAAGTATATT | 480  |
| GTTGTTATTA | TTAAGGTGCG | AGATGAGATA | AGGTCTACAT | GGACAGCACA | AAACCCACCC | 540  |
| CTAATGCGAA | TAGGGGTGGG | TTTTTTTCGT | TCGTTGCGAA | TACGAACGTG | TGGGTTAGAG | 600  |
| ACAACCTGCG | AGATTATCGT | CTAATCATCT | AACCAATGAT | CCACTAGTAT | TAATACTAGT | 660  |
| CCCACAAAAA | GTGGAGCAAT | AACCAATGAG | ATAAGGTTTT | CCATAAACAG | CACCCCCTTT | 720  |
| CAGGGGCAAG | TTGCCACTTA | CTAATATAGC | ACAGCTCCTT | TATTGTTCTT | AGTCTAAATC | 780  |
| TGATAAATCT | TTTCTTGTTT | AAAAATATAG | ACCACTTAAA | AGCTTATAAC | GGTACTAGAT | 840  |
| TTTTCAGATA | CCCCAATTAC | CTACTTAAAA | CGTCTCTCTT | TTTCGTTTTA | AGATGTTTAA | 900  |
| AATTATTTTC | TATGAATTAT | ACACAAATGT | GCTTAAATCG | TCTTAAATCG | TCTTAAATG  | 960  |
| TGGTCTGTGT | TGAGAATACA | ACGACTTTGT | TTGGTCGTAC | CTCTAAATCT | GTTTGCTGTG | 1020 |
| AACGAGGGTA | GCGAAGTGAA | CTTTTTGTTG | CTAACGCTCT | TGGTTTTGTC | TTTTGATTTT | 1080 |
| ATAAAATGTG | GATGTAATCC | ACTCCTTACT | AGGGGTTTAA | TCTTTATAAA | ATAAAGGAGC | 1140 |
| TTGCGAATGC | AAGGTGCCCT | TTTTTCTTTG | TCTGACTACT | AGGGACAAAT | TATCTGAGTA | 1200 |
| TGAACAAGAT | TTTGTCTGTT | CTTGCGCGTA | TTTATTAATA | TATATTTTAA | GAGATATTTT | 1260 |
| AAGAGATATT | TTAAACCTT  | TTTAGGGGTG | AGCTCAGCCT | TAGAGAGAGT | AAGCATTGAA | 1320 |



PATENT  
Atty Docket 07254/051001

|    |                                                                    |      |
|----|--------------------------------------------------------------------|------|
|    | GCATAGTACT AGGGACAAAT TATCTGACTA CTAGGGATAA ATTATCTGAC TACTAGGGAC  | 1380 |
|    | AAATTATCTG ACTACTAGGG ATAAATTATC TGACTIONTAG GGACGCACTT TACTTTGTGT | 1440 |
| 5  | ATCGTATCGT TTATAATCTT TATATGTGAG GGGAGGTCGA AAGGATTGGA AAAGAAAACG  | 1500 |
|    | AATTTAAAAA TTGCATATCA AAATGAATTG AATCTGGTTC CACTTAAAAA TTTCAATGCT  | 1560 |
|    | AAACAAATGG ATTTATTCTT TGCTTTGTGT GCCCGAATGA AAAATAAAGG GCTTAGAAAG  | 1620 |
| 10 | GTATCTTTTA CGTTTGAAGA ACTAAGAGAA CTAAGTGATT ACAAGATAAC TGCTATCGAA  | 1680 |
|    | CCGTTACAGA ATGATTTAGA ACAATTATAC AAAAAATGT TAAACCTAAC ATACAGAACG   | 1740 |
| 15 | GAGACAGAAA CAAAAATCAG TTATTTTCGTT TTATTTACTG GGTTTGTGAT TGATAAATCA | 1800 |
|    | GAGCAAATTG TTGAAGTTAG TGTAACCCCA GACTTAGAAC ATATCATTAA CGGTATCTCT  | 1860 |
|    | AGTGAGTTCA ATAAATTTGA GTTACTAGCA TTCACAAGTA TCCAGTCGAA ATATACGAAA  | 1920 |
| 20 | ACACTCTTTA GATTGCTTAT GCAGTTTCAA TCAACTGGGT TTTATGTGGT TAAAATTGAA  | 1980 |
|    | GATTTACAGAG CACTTTTAGA CATTCCAAAA TCTTATCAAA TGACTGACAT AACCCAACGG | 2040 |
| 25 | ATATTGAAAC CTAGTTTAAAT TGAGTTAAGT CAGTACTTTA ATGATTTAAA AGTTAATAAA | 2100 |
|    | ATTAAAGCTC GAAAGGGTAA TAAAATAGAC CGTTTAGAAT TCACTTTCTC CGGTCTAAAG  | 2160 |
|    | ACTGATTTAC CTAAAGTTCC ATTGCACGAC TGGACGAAAT AAAAAAGGA CCTCCCCCTC   | 2220 |
| 30 | ACATTTAAGC AAGTAGGAAC GTCCCTCGCA ATCCACGAAG ACTGCTGATT CATTTTAGCA  | 2280 |
|    | TATATTGTGC GGGACTTCTA AATAAATTAT ATTTGGAGGT CATTTTTATG TCGAATAAGT  | 2340 |
| 35 | ACTTGAAAAA AAGAAAGCGT CAAGCTAAGC AGGTAGCTGA TTTGTACGAT TTAATTATTG  | 2400 |
|    | GGGTTGAACA TGCTGGCAGC TCGTTAATTG CGTTGTATGA GGAATTAAA CCTCTCAAT    | 2460 |
|    | ATCGAATTTT TATTCTTTTG TCTTATTCTA GTTTTGAAAA TAAATTAAAC TTATACAACA  | 2520 |
| 40 | AAGCGATTTT AAGAACTGAA GTTTATTCTT TAGAAAAAAA ATTAAACGAA AAAATAAATG  | 2580 |
|    | CTCAATCAG AATTGCGCAA AAAAATAAAA AAGAAATTGC GGTAATTGAT TTCACAAAAC   | 2640 |
| 45 | AAAAAGAAAA ACTCAAAAGA GAATTACTTA GTTTTGAAAA TGATAAAGAA ATGAACTTA   | 2700 |
|    | TGGATTGCGA ATTAACAA TTTTCATGAAA ATAAACGTT AGCTGATATT AATGATCAGT    | 2760 |
|    | TTTTTATGAC GGTACAAAAT AGTTTAATTT TGCTGCATAA AAAAGCACCT TTAACATTAA  | 2820 |
| 50 | AATTAATTTG TTTGAAAAAT TATATTCGCC TTTGCAAAAA TTATTTTCTA AAGAATATAT  | 2880 |
|    | TTAATGTTT TTTGAAAAAA ATAGTAACAT GGGAACATGT TGCTCTGCTC GCAAAAGGAA   | 2940 |

AAATATTTAA ACTAATAAAA AACCGTCGGA GACCAGCCAA CCAATAGGTT GGCTTTAAGT 3000  
TTAAGCCTAC GTTGACAACT GTCAATGTAT AAGTGCGCCC TTTGGGTGTT TTATTTTTTG 3060  
5 TTTAACTATT ATTTTCTGCA TAGGTTTTTT ATTTTATTATA ATTTGATTTT CAAGAAAGGG 3120  
ATGAACCTAA AATGATTTAT AAACAAAAAA AGAAAGAAGA TGTTTTTGGA TTTCCTAAAG 3180  
TTTAAACAAT TGCTGATTTG AGTACGAGAT GGAAAATGTC ACGTCAAGCT ATCCATAAAA 3240  
10 AAATTCAAGA AGATTTATTA TTTCCTATGC CTGTTCAAAT TGTCTCAAAT GGAAAATTA 3300  
AATTGTTTTT ATTTGTTGAT ATTGAAAAAT ACGAAAAAAA TCGTCCGTGG TTATTAGACA 3360  
15 TTAATTATCG AAATGAACGA CAACTTTGGA TTTACAAAAA TGGTTTTTTT AAATAGCAAG 3420  
TTAGTCAATT ACCTTATACC TTGTTGGATA TCTTTGGATA AAAAAATAGT TGTAT 3475

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 228 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: Divergicin structural gene

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATG AAA AAA CAA ATT TTA AAA GGG TTG GTT ATA GTT GTT TGT TTA TCT 48  
45 Met Lys Lys Gln Ile Leu Lys Gly Leu Val Ile Val Val Cys Leu Ser  
1 5 10 15  
GGG GCA ACA TTT TTC TCA ACA CCA CAA CAA GCT TCT GCT GCT GCA CCG 96  
50 Gly Ala Thr Phe Phe Ser Thr Pro Gln Gln Ala Ser Ala Ala Ala Pro  
20 25 30  
AAA ATT ACT CAA AAA CAA AAA AAT TGT GTT AAT GGA CAA TTA GGT GGA 144  
Lys Ile Thr Gln Lys Gln Lys Asn Cys Val Asn Gly Gln Leu Gly Gly  
35 40 45

ATG CTT GCT GGA GCT TTG GGT GGA CCT GGC GGA GTT GTG TTA GGT GGT 192  
Met Leu Ala Gly Ala Leu Gly Gly Pro Gly Gly Val Val Leu Gly Gly  
50 55 60

5 ATA GGT GGT GCA ATA GCA GGA GGT TGT TTT AAT TA 228  
Ile Gly Gly Ala Ile Ala Gly Gly Cys Phe Asn  
65 70 75

10 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 amino acids  
(B) TYPE: amino acid  
15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

20 Met Lys Lys Gln Ile Leu Lys Gly Leu Val Ile Val Val Cys Leu Ser  
1 5 10 15  
25 Gly Ala Thr Phe Phe Ser Thr Pro Gln Gln Ala Ser Ala Ala Ala Pro  
20 25 30  
Lys Ile Thr Gln Lys Gln Lys Asn Cys Val Asn Gly Gln Leu Gly Gly  
35 40 45  
30 Met Leu Ala Gly Ala Leu Gly Gly Pro Gly Gly Val Val Leu Gly Gly  
50 55 60  
35 Ile Gly Gly Ala Ile Ala Gly Gly Cys Phe Asn  
65 70 75

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 171 base pairs  
(B) TYPE: nucleic acid  
40 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (vii) IMMEDIATE SOURCE:

(B) CLONE: Divergicin immunity genes

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..171

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATG AAA ATA AAA TGG TAC TGG GAA TCT CTG ATT GAA ACC TTA ATA TTT 48  
Met Lys Ile Lys Trp Tyr Trp Glu Ser Leu Ile Glu Thr Leu Ile Phe 15  
1 5 10  
ATA ATT GTT CTT CTT GTA TTT TTT TAT AGA AGT TCT GGT TTT TCT TTA 96  
Ile Ile Val Leu Leu Val Phe Phe Tyr Arg Ser Ser Gly Phe Ser Leu 30  
20 25  
AAA AAT TTA GTT TTA GGA AGT TTA TTT TAT TTG ATA GCA ATT GGT CTT 144  
Lys Asn Leu Val Leu Gly Ser Leu Phe Tyr Leu Ile Ala Ile Gly Leu 45  
35 40  
TTT AAT TAT AAA AAG ATA AAC AAA TA 171  
Phe Asn Tyr Lys Lys Ile Asn Lys 55  
50

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 56 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Lys Ile Lys Trp Tyr Trp Glu Ser Leu Ile Glu Thr Leu Ile Phe 15  
1 5 10  
Ile Ile Val Leu Leu Val Phe Phe Tyr Arg Ser Ser Gly Phe Ser Leu 30  
20 25  
Lys Asn Leu Val Leu Gly Ser Leu Phe Tyr Leu Ile Ala Ile Gly Leu 45  
35 40  
Phe Asn Tyr Lys Lys Ile Asn Lys 55  
50

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 124 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: Divergicin signal sequence

(ix) FEATURE:  
(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 1..123

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..123

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATC | TTG | GTA | TCA | CAA | ACT | AAT | TTG | GAG | GTT | GGT | ATA | TAT | GAA | AAA | ACA | 48  |
| Ile | Leu | Val | Ser | Gln | Thr | Asn | Leu | Glu | Val | Gly | Ile | Tyr | Glu | Lys | Thr |     |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 96  |
| AAT | TTT | AAA | AGG | GTT | GGT | TAT | AGT | TGT | TTG | TTT | ATC | TGG | GGC | AAC | ATT |     |
| Asn | Phe | Lys | Arg | Val | Gly | Tyr | Ser | Cys | Leu | Phe | Ile | Trp | Gly | Asn | Ile |     |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     | 30  |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 124 |
| TTT | CTC | AAC | ACC | ACA | ACA | AGC | TTC | TGC | T   |     |     |     |     |     |     |     |
| Phe | Leu | Asn | Thr | Thr | Thr | Ser | Phe | Cys |     |     |     |     |     |     |     |     |
|     |     | 35  |     |     |     | 40  |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 41 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |    |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| Ile | Leu | Val | Ser | Gln | Thr | Asn | Leu | Glu | Val | Gly | Ile | Tyr | Glu | Lys | Thr |    |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |     |    |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 50 |
| Asn | Phe | Lys | Arg | Val | Gly | Tyr | Ser | Cys | Leu | Phe | Ile | Trp | Gly | Asn | Ile |    |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     | 30  |     |     |     |    |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |    |
| Phe | Leu | Asn | Thr | Thr | Thr | Ser | Phe | Cys |     |     |     |     |     |     |     |    |

35

40

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 675 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: Brochocin-C

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| GCTATTTTGA GAAATATTAA CCAATAGTAA AAATTATCAT GCTATCTTTT GTATGTAATA  | 60  |
| AAAATTATTT AAAGGAGGGT GTTTCATCAT GCACAAGGTA AAAAAATTAA ACAATCAAGA  | 120 |
| GTTACAACAG ATCGTGGGAG GTTACAGTTC AAAAGATTGT CTAAAAGATA TTGGTAAAGG  | 180 |
| AATTGGTGCT GGTACAGTAG CTGGGGCAGC CGGCGGTGGC CTAGCTGCAG GATTAGGTGC  | 240 |
| TATCCCAGGA GCATTCTGTTG GAGCACATTT TGGAGTAATC GGCGGATCTG CCGCATGCAT | 300 |
| TGGTGGATTA TTAGGTAAGT AGGAGGTTAT ATTTATGAAA AAAGAACTAT TGAATAAAAA  | 360 |
| TGAAATGAGT AGAATTATCG GCGGCAAAAT AAATTGGGGA AATGTTGGCG GTTCTTGTGT  | 420 |
| TGGAGGTGCA GTAATTGGAG GCGCCCTCGG TGGACTAGGT GGAGCTGGCG GAGGTTGCAT  | 480 |
| TACAGGAGCT ATCGGAAGTA TTTGGGATCA ATGGTAAAAA CTATACTATT TTCGGTTGTA  | 540 |
| ATTTCAATCG TTGCATTATG TAACTTTTAA ATAAAAAAG ATGTGTCTTC AAAAAAAAAA   | 600 |
| TTATTTTAA CAGGTTCTAT TGCTGTCTTT CTAATTATCT ATGATTTTCT ATGGATTATA   | 660 |
| TTCTCTAACT AGTAC                                                   | 675 |

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 234 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vii) IMMEDIATE SOURCE:

(B) CLONE: Brochocin-C peptide A

(ix) FEATURE:

(A) NAME/KEY: CDS

15 (B) LOCATION: 1..234

(ix) FEATURE:

(A) NAME/KEY: sig\_peptide

(B) LOCATION: 1..54

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

25 ATG CAC AAG GTA AAA AAA TTA AAC AAT CAA GAG TTA CAA CAG ATC GTG 48  
Met His Lys Val Lys Lys Leu Asn Asn Gln Glu Leu Gln Gln Ile Val  
1 5 10 15  
GGA GGT TAC AGT TCA AAA GAT TGT CTA AAA GAT ATT GGT AAA GGA ATT 96  
Gly Gly Tyr Ser Ser Lys Asp Cys Leu Lys Asp Ile Gly Lys Gly Ile  
20 25 30  
GGT GCT GGT ACA GTA GCT GGG GCA GCC GGC GGT GGC CTA GCT GCA GGA 144  
Gly Ala Gly Thr Val Ala Gly Ala Ala Gly Gly Gly Leu Ala Ala Gly  
35 40 45  
TTA GGT GCT ATC CCA GGA GCA TTC GTT GGA GCA CAT TTT GGA GTA ATC 192  
Leu Gly Ala Ile Pro Gly Ala Phe Val Gly Ala His Phe Gly Val Ile  
50 55 60  
40 GGC GGA TCT GCC GCA TGC ATT GGT GGA TTA TTA GGT AAC TA 234  
Gly Gly Ser Ala Ala Cys Ile Gly Gly Leu Leu Gly Asn  
65 70 75

45 (2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met His Lys Val Lys Lys Leu Asn Asn Gln Glu Leu Gln Gln Ile Val  
1 5 10 15  
5 Gly Gly Tyr Ser Ser Lys Asp Cys Leu Lys Asp Ile Gly Lys Gly Ile  
20 25 30  
Gly Ala Gly Thr Val Ala Gly Ala Ala Gly Gly Gly Leu Ala Ala Gly  
35 40 45  
10 Leu Gly Ala Ile Pro Gly Ala Phe Val Gly Ala His Phe Gly Val Ile  
50 55 60  
Gly Gly Ser Ala Ala Cys Ile Gly Gly Leu Leu Gly Asn  
65 70 75  
15

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: Brochocin-C peptide B

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..183

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide  
(B) LOCATION: 1..51

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

45 ATG AAA AAA GAA CTA TTG AAT AAA AAT GAA ATG AGT AGA ATT ATC GGC 48  
Met Lys Lys Glu Leu Leu Asn Lys Asn Glu Met Ser Arg Ile Ile Gly  
1 5 10 15  
50 GGC AAA ATA AAT TGG GGA AAT GTT GGC GGT TCT TGT GTT GGA GGT GCA 96  
Gly Lys Ile Asn Trp Gly Asn Val Gly Gly Ser Cys Val Gly Gly Ala  
20 25 30  
GTA ATT GGA GGC GCC CTC GGT GGA CTA GGT GGA GCT GGC GGA GGT TGC 144



Val Ile Gly Gly Ala Leu Gly Gly Leu Gly Gly Ala Gly Gly Gly Cys  
35 40 45

5 ATT ACA GGA GCT ATC GGA AGT ATT TGG GAT CAA TGG TA  
Ile Thr Gly Ala Ile Gly Ser Ile Trp Asp Gln Trp  
50 55 60

183

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

20 Met Lys Lys Glu Leu Leu Asn Lys Asn Glu Met Ser Arg Ile Ile Gly  
1 5 10 15  
Gly Lys Ile Asn Trp Gly Asn Val Gly Gly Ser Cys Val Gly Gly Ala  
20 25 30  
25 Val Ile Gly Gly Ala Leu Gly Gly Leu Gly Gly Ala Gly Gly Gly Cys  
35 40 45  
30 Ile Thr Gly Ala Ile Gly Ser Ile Trp Asp Gln Trp  
50 55 60

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 162 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: Brochocin-C immunity peptide

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

5     ATG GTA AAA ACT ATA CTA TTT TCG GTT GTA ATT TCA TTC GTT GCA TTA     48  
      Met Val Lys Thr Ile Leu Phe Ser Val Val Ile Ser Phe Val Ala Leu  
      1                     5                     10                     15

10     TGT AAC TTT TTA ATA AAA AAA GAT GTG TCT TCA AAA AAA AAA TTA TTT     96  
      Cys Asn Phe Leu Ile Lys Lys Asp Val Ser Ser Lys Lys Lys Leu Phe  
                  20                     25                     30

15     TTA ACA GGT TCT ATT GCT GTC TTT CTA ATT ATC TAT GAT TTT CTA TGG     144  
      Leu Thr Gly Ser Ile Ala Val Phe Leu Ile Ile Tyr Asp Phe Leu Trp  
                  35                     40                     45

15     ATT ATA TTC TCT AAC TA     162  
      Ile Ile Phe Ser Asn  
                  50

20 (2) INFORMATION FOR SEQ ID NO:27:

- 25     (i) SEQUENCE CHARACTERISTICS:  
      (A) LENGTH: 53 amino acids  
      (B) TYPE: amino acid  
      (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

30     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

30     Met Val Lys Thr Ile Leu Phe Ser Val Val Ile Ser Phe Val Ala Leu  
      1                     5                     10                     15

35     Cys Asn Phe Leu Ile Lys Lys Asp Val Ser Ser Lys Lys Lys Leu Phe  
                  20                     25                     30

40     Leu Thr Gly Ser Ile Ala Val Phe Leu Ile Ile Tyr Asp Phe Leu Trp  
                  35                     40                     45

40     Ile Ile Phe Ser Asn  
                  50

45 (2) INFORMATION FOR SEQ ID NO:28:

- 50     (i) SEQUENCE CHARACTERISTICS:  
      (A) LENGTH: 2226 base pairs  
      (B) TYPE: nucleic acid  
      (C) STRANDEDNESS: single  
      (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: DNA (genomic)

      (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: Enterocin 900 operon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

10 AAGCTTTACT TGATATTAGT TCTGAGTTCT GCCTGATTTA TCAGTAACAT AACTCTAGAG 60  
ATAACTGCGT CGCTATCTCA AGTTCTTTTT TCTTTTCTTA CAAAATAAAT ATACTTTATT 120  
TCATTTTATA AGTCAACGTT TTCATTGCTT ATAAATTAGT TTTTTTAATC ATCTCAGTAA 180  
15 TAATTTGCTA TGTCAGTTCTG ATCAATACCA TTTGCATGAA AGTACAGCTA TAAGCCAATC 240  
ACACCTAACC TACTCTTAAT CGTATAATGA TTCCAGTTAG CAAGGTCTTT AATACTCATG 300  
20 ATCCCATCT GGTGAGTTT TTTTCATTCG AAACCCTATG CCCCAAAAGA CCGTCATTTT 360  
AGGAATATTC CATACTTTCT CAGGAACATT TTGATACGTC CATTGAGCAA TAAACCCTTC 420  
ATTGTGCTTC GCTTCATTGT CTAAGGCGAG TTTGGCCAAA GGGGATTATC TCCGCGACAC 480  
25 CTACCGTAGC AATCAATCCT AATTCTTCTT TAATACGTTT CTTGGATCAT TTGAACGAAT 540  
TTTTTCCTTC TCTGACTTCT TGTTCTTCA GTCGTAAAAA TATTCAGTGA TCTGGTCACT 600  
30 TTTAAAATGG GTTCATCGAT TGGATACATC AGTAGATCTT CGTCAGCCAC ATATCTTTTG 660  
AAAATATTAT TTACCCGCAT ATTCGCTTG ATATATAGGT TCATACGTGG TGGAACAACG 720  
TATGATGTTT TAGGAAATAG TTGTGATAAA TCACGTGGTC TACTCACATT TGTAATATCA 780  
35 TACCGCTTTT TTGCTTCAGG AGAAGAAGCT CTAATATCAA TCCTAAACCA GTATTGTCAG 840  
CGCGACTCAT AACAACAAGT TCTGTTGTTA ATGGATCAAA ATTTCTTTCT ATACACTCGA 900  
40 TACTCGCATA AAAAGGCTTC ATGTCGATTA GAAAATAATC ATTTACTGAT TCTTTCGAAT 960  
AATCCAGCAT GAATAACACC CATTCTTTTT CACATTACAC AAACGTAAGT TAGGAAATAT 1020  
AAAGAAGAAA ACTAAATAGC ACTAAACAAA CAAGACAACT CATGCTTATT CCGTATAAGA 1080  
45 AACTACATAT TATGTAACT AGTTATTAAA ATAACATATT TAATAAAATT AAATTGTGAT 1140  
TTTATAGGTT TCAGGAATGA AAAAGCCTTA TTTCAGGAAG TTTTAACTG TTTGCTATAG 1200  
50 ATGTATGTCA TGATAGCATC GTAATAAAAA TACTCTAAAA GGAGCGAGTT TAAATATGCA 1260  
AAATGTAAAA GAATTAAGTA CGAAAGAGAT GAAACAAATT ATCGGTGGAG AAAATGATCA 1320  
CAGAATGCCT AATGAGTTAA ATAGACCTAA CAACTTATCT AAAGGTGGAG CAAAATGTGG 1380

TGCTGCAATT GCTGGGGGAT TATTTGGAAT CCCAAAAGGA CCACTAGCAT GGGCTGCTGG 1440  
GTTAGCAAT GTATACTCTA AATGCAACTA AAAAAGAAGA GAAAAAACTC ATTACGAGTT 1500  
5 TTTTCTCTTC TTTTTTTGCA TGAAATTAGG AATAACTAAT AAAACAATAG CAATCAATAG 1560  
TAAATCTTA CTTAATATAG TTTCGGAAAA AATAAATAAT CCTAAATTTA TAATTACTGC 1620  
TAAAAAATG CATAAATTAT ACTCTAAATT ATTTTTTTTT AAATTCATAA TATAAACATC 1680  
10 CTCTCTTTAA TTAGTCTACC ATTCCGAAAT ATTTTCATCCC CAGCTCTTTT TTTACTAATA 1740  
TACCAACTAC ATTTAATAAC AAAATAACTA GTAAACTTAA TATTTTTTAGT GGCATAGAAT 1800  
15 ATTCAAAAAT AAATAAAGGC ACCATACATG TAGCTATCAA TATAAATACA GAACTTACGT 1860  
ATTTTATTAT TTTACGGAAC ATTATAACCT ATTACAACTC CGCAAATAGC CATAGCCCAT 1920  
ACCATAGATA AGATTTTTTAC CAGCACCACC ACCACATGTT TGTTTTATCT CTTTCATACT 1980  
20 TAATTTTTTT ACATTTTGCA TGTCTCTACA TGCTCCTTTT AAAGTTTTTT TAGAACCTCA 2040  
CGACTATAAC ATGGATAATT TAATCGTGGT CAAAACTTC CTGAAATAGG GTGTTTCATA 2100  
25 TCCTGAACAC GAATTTTTAG TCAATTTTCG AAAAATGAAA CTTTAAAATT TCTTTGACCA 2160  
GAACTCTATT TATTCTGTG TTGTTCTTC GAATAGGTC CCGTATATCT TTTTATTTG 2220  
AAGCTT 2226

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 216 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: Enterocin 900 peptide

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..216

(ix) FEATURE:  
(A) NAME/KEY: sig\_peptide

(B) LOCATION: 1..54

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

5  
ATG CAA AAT GTA AAA GAA TTA AGT ACG AAA GAG ATG AAA CAA ATT ATC 48  
Met Gln Asn Val Lys Glu Leu Ser Thr Lys Glu Met Lys Gln Ile Ile  
1 5 10 15  
10 GGT GGA GAA AAT GAT CAC AGA ATG CCT AAT GAG TTA AAT AGA CCT AAC 96  
Gly Gly Glu Asn Asp His Arg Met Pro Asn Glu Leu Asn Arg Pro Asn  
20 25 30  
15 AAC TTA TCT AAA GGT GGA GCA AAA TGT GGT GCT GCA ATT GCT GGG GGA 144  
Asn Leu Ser Lys Gly Gly Ala Lys Cys Gly Ala Ala Ile Ala Gly Gly  
35 40 45  
TTA TTT GGA ATC CCA AAA GGA CCA CTA GCA TGG GCT GCT GGG TTA GCA 192  
Leu Phe Gly Ile Pro Lys Gly Pro Leu Ala Trp Ala Ala Gly Leu Ala  
50 55 60  
AAT GTA TAC TCT AAA TGC AAC TA 216  
Asn Val Tyr Ser Lys Cys Asn  
65 70

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

40 Met Gln Asn Val Lys Glu Leu Ser Thr Lys Glu Met Lys Gln Ile Ile 15  
1 5 10  
Gly Gly Glu Asn Asp His Arg Met Pro Asn Glu Leu Asn Arg Pro Asn 30  
20 25 30  
45 Asn Leu Ser Lys Gly Gly Ala Lys Cys Gly Ala Ala Ile Ala Gly Gly 45  
35 40  
Leu Phe Gly Ile Pro Lys Gly Pro Leu Ala Trp Ala Ala Gly Leu Ala 60  
50 55 60  
50 Asn Val Tyr Ser Lys Cys Asn 70  
65 70

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: Colicin V prepeptide

(ix) FEATURE:

- (A) NAME/KEY: Disulfide-bond  
(B) LOCATION: 91..102

(ix) FEATURE:

- (A) NAME/KEY: Cleavage-site  
(B) LOCATION: 15..16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Thr | Leu | Thr | Leu | Asn | Glu | Leu | Asp | Ser | Val | Ser | Gly | Gly | Ala |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Ser | Gly | Arg | Asp | Ile | Ala | Met | Ala | Ile | Gly | Thr | Leu | Ser | Gly | Gln | Phe |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Val | Ala | Gly | Gly | Ile | Gly | Ala | Ala | Ala | Gly | Gly | Val | Ala | Gly | Gly | Ala |
|     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |
| Ile | Tyr | Asp | Tyr | Ala | Ser | Thr | His | Lys | Pro | Asn | Pro | Ala | Met | Ser | Pro |
|     |     | 50  |     |     |     | 55  |     |     |     | 60  |     |     |     |     |     |
| Ser | Gly | Leu | Gly | Gly | Thr | Ile | Lys | Gln | Lys | Pro | Glu | Gly | Ile | Pro | Ser |
|     |     | 65  |     |     |     | 70  |     |     |     | 75  |     |     |     | 80  |     |
| Glu | Ala | Trp | Asn | Tyr | Ala | Ala | Gly | Arg | Leu | Cys | Asn | Trp | Ser | Pro | Asn |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Asn | Leu | Ser | Asp | Val | Cys | Leu |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     | 100 |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vii) IMMEDIATE SOURCE:  
(B) CLONE: Colicin V

(ix) FEATURE:

15 (A) NAME/KEY: Disulfide-bond

(B) LOCATION: 76..87

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

20 Ala Ser Gly Arg Asp Ile Ala Met Ala Ile Gly Thr Leu Ser Gly Gln  
1 5 10 15

Phe Val Ala Gly Gly Ile Gly Ala Ala Ala Gly Gly Val Ala Gly Gly  
20 25 30

25 Ala Ile Tyr Asp Tyr Ala Ser Thr His Lys Pro Asn Pro Ala Met Ser  
35 40 45

30 Pro Ser Gly Leu Gly Gly Thr Ile Lys Gln Lys Pro Glu Gly Ile Pro  
50 55 60

Ser Glu Ala Trp Asn Tyr Ala Ala Gly Arg Leu Cys Asn Trp Ser Pro  
65 70 75 80

35 Asn Asn Leu Ser Asp Val Cys Leu  
85

(2) INFORMATION FOR SEQ ID NO:33:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 675 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: carnobacteriocin BM1

(ix) FEATURE:  
 (A) NAME/KEY: RBS  
 (B) LOCATION: 89..93

5 (ix) FEATURE:  
 (A) NAME/KEY: RBS  
 (B) LOCATION: 290..293

10 (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 103..285

15 (ix) FEATURE:  
 (A) NAME/KEY: sig\_peptide  
 (B) LOCATION: 103..150

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 303..566

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| TCGAGATACG TTTATCCATG GTTCAGGATG ATTTTATCAA CGAGTTAATT ATTTATGCTA | 60  |
| CAGTAAACTT GTTACTAAAT ACTTTATAAG GAGTGTATGT AC ATG AAA AGC GTT    | 114 |
| Met Lys Ser Val                                                   |     |
| 1                                                                 |     |
| AAA GAA CTA AAT AAA AAA GAA ATG CAA CAA ATT AAT GGT GGA GCT ATC   | 162 |
| Lys Glu Leu Asn Lys Lys Glu Met Gln Gln Ile Asn Gly Gly Ala Ile   |     |
| 5 10 15 20                                                        |     |
| TCT TAT GGC AAT GGT GTT TAT TGT AAC AAA GAG AAA TGT TGG GTA AAC   | 210 |
| Ser Tyr Gly Asn Gly Val Tyr Cys Asn Lys Glu Lys Cys Trp Val Asn   |     |
| 25 30 35                                                          |     |
| AAG GCA GAA AAC AAA CAA GCT ATT ACT GGA ATA GTT ATC GGT GGA TGG   | 258 |
| Lys Ala Glu Asn Lys Gln Ala Ile Thr Gly Ile Val Ile Gly Gly Trp   |     |
| 40 45 50                                                          |     |
| GCT TCT AGT TTA GCA GGA ATG GGA CAT TAAAGAGGTA TCTAGTT ATG ATA    | 308 |
| Ala Ser Ser Leu Ala Gly Met Gly His                               |     |
| 55 60 1                                                           |     |
| AAA GAT GAA AAA ATA AAT AAA ATC TAT GCT TTA GTT AAG AGC GCA CTT   | 356 |
| Lys Asp Glu Lys Ile Asn Lys Ile Tyr Ala Leu Val Lys Ser Ala Leu   |     |
| 5 10 15                                                           |     |
| GAT AAT ACG GAT GTT AAG AAT GAT AAA AAA CTT TCT TTA CTT CTT ATG   | 404 |
| Asp Asn Thr Asp Val Lys Asn Asp Lys Lys Leu Ser Leu Leu Leu Met   |     |
| 20 25 30                                                          |     |
| AGA ATA CAA GAA ACA TCA ATT AAT GGA GAA CTA TTT TAC GAT TAT AAA   | 452 |



Arg Ile Gln Glu Thr Ser Ile Asn Gly Glu Leu Phe Tyr Asp Tyr Lys  
35 40 45 50

5 AAA GAA TTA CAG CCA GCT ATT AGT ATG TAC TCT ATT CAA CAT AAC TTT 500  
Lys Glu Leu Gln Pro Ala Ile Ser Met Tyr Ser Ile Gln His Asn Phe  
55 60 65

10 CGG GTT CCT GAC GAT CTA GTA AAA CTG TTA GCA TTA GTT CAA ACA CCT 548  
Arg Val Pro Asp Asp Leu Val Lys Leu Leu Ala Leu Val Gln Thr Pro  
70 75 80

15 AAA GCT TGG TCA GGG TTT TAACTTTAGT TCCAGATGAG TTAAAATCCT 596  
Lys Ala Trp Ser Gly Phe  
85

TAAAAATAAG GAATAATGGT AAATCAGCAT TCCTTATTTT TATAGTCATC ACACTATAAC 656

TTTACTTAAA GATGTTCGA 675

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Ser Val Lys Glu Leu Asn Lys Lys Glu Met Gln Gln Ile Asn  
1 5 10 15

35 Gly Gly Ala Ile Ser Tyr Gly Asn Gly Val Tyr Cys Asn Lys Glu Lys  
20 25 30

40 Cys Trp Val Asn Lys Ala Glu Asn Lys Gln Ala Ile Thr Gly Ile Val  
35 40 45

Ile Gly Gly Trp Ala Ser Ser Leu Ala Gly Met Gly His  
50 55 60

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Ile Lys Asp Glu Lys Ile Asn Lys Ile Tyr Ala Leu Val Lys Ser  
1 5 10 15  
Ala Leu Asp Asn Thr Asp Val Lys Asn Asp Lys Lys Leu Ser Leu Leu  
5 20 25 30  
Leu Met Arg Ile Gln Glu Thr Ser Ile Asn Gly Glu Leu Phe Tyr Asp  
35 40 45  
10 Tyr Lys Lys Glu Leu Gln Pro Ala Ile Ser Met Tyr Ser Ile Gln His  
50 55 60  
Asn Phe Arg Val Pro Asp Asp Leu Val Lys Leu Leu Ala Leu Val Gln  
15 65 70 75 80  
Thr Pro Lys Ala Trp Ser Gly Phe  
85

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1907 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: carnobacteriocin B2 operon

(ix) FEATURE:  
(A) NAME/KEY: -35\_signal  
(B) LOCATION: 165..170

(ix) FEATURE:  
(A) NAME/KEY: -10\_signal  
(B) LOCATION: 189..193

(ix) FEATURE:  
(A) NAME/KEY: RBS  
(B) LOCATION: 221..225

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 236..433

(ix) FEATURE:  
(A) NAME/KEY: RBS

(B) LOCATION: 460..465

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 473..805

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1025..1267

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

|    |                                                                    |     |
|----|--------------------------------------------------------------------|-----|
| 15 | AAGCTTTTAT AGTACAATTA TTTATGCGTG CTATGCAATA GCTATTGTAT ATACTATTTT  | 60  |
|    | TACTATGAGA AAAGATTCTT ATGAAAATAA CAAAAATAAT CGTAAAAAAG TTATATAGCA  | 120 |
|    | TTTATTTTCAT TTATGAATTC AAATACCCTG GTTCAAGATG TATTTTCCAA AAAAATGTTC | 180 |
| 20 | AGATATGATA TAGTTTTTTT GAAATACAAA TATAAAATAA AGGAGTTTGA TTTAG ATG   | 238 |
|    | Met                                                                |     |
|    | 1                                                                  |     |
| 25 | AAT AGC GTA AAA GAA TTA AAC GTG AAA GAA ATG AAA CAA TTA CAC GGT    | 286 |
|    | Asn Ser Val Lys Glu Leu Asn Val Lys Glu Met Lys Gln Leu His Gly    |     |
|    | 5 10 15                                                            |     |
| 30 | GGA GTA AAT TAT GGT AAT GGT GTT TCT TGC AGT AAA ACA AAA TGT TCA    | 334 |
|    | Gly Val Asn Tyr Gly Asn Gly Val Ser Cys Ser Lys Thr Lys Cys Ser    |     |
|    | 20 25 30                                                           |     |
| 35 | GTT AAC TGG GGA CAA GCC TTT CAA GAA AGA TAC ACA GCT GGA ATT AAC    | 382 |
|    | Val Asn Trp Gly Gln Ala Phe Gln Glu Arg Tyr Thr Ala Gly Ile Asn    |     |
|    | 35 40 45                                                           |     |
| 40 | TCA TTT GTA AGT GGA GTC GCT TCT GGG GCA GGA TCC ATT GGT AGG AGA    | 430 |
|    | Ser Phe Val Ser Gly Val Ala Ser Gly Ala Gly Ser Ile Gly Arg Arg    |     |
|    | 50 55 60 65                                                        |     |
| 40 | CCG TAAATATATA AATACAATAT AGAGCAAGGT GGTGATACA ATG GAT ATA AAG     | 484 |
|    | Pro Met Asp Ile Lys                                                |     |
|    | 1                                                                  |     |
| 45 | TCT CAA ACA TTA TAT TTG AAT CTA AGC GAG GCA TAT AAA GAC CCT GAA    | 532 |
|    | Ser Gln Thr Leu Tyr Leu Asn Leu Ser Glu Ala Tyr Lys Asp Pro Glu    |     |
|    | 5 10 15 20                                                         |     |
| 50 | GTA AAA GCT AAT GAA TTC TTA TCA AAA TTA GTT GTA CAA TGT GCT GGG    | 580 |
|    | Val Lys Ala Asn Glu Phe Leu Ser Lys Leu Val Val Gln Cys Ala Gly    |     |
|    | 25 30 35                                                           |     |
|    | AAA TTA ACA GCT TCA AAC AGT GAG AAC AGT TAT ATT GAA GTA ATA TCA    | 628 |
|    | Lys Leu Thr Ala Ser Asn Ser Glu Asn Ser Tyr Ile Glu Val Ile Ser    |     |
|    | 40 45 50                                                           |     |

|    |                                                                   |      |
|----|-------------------------------------------------------------------|------|
|    | TTG CTA TCT AGG GGT ATT TCT AGT TAT TAT TTA TCC CAT AAA CGT ATA   | 676  |
|    | Leu Leu Ser Arg Gly Ile Ser Ser Tyr Tyr Leu Ser His Lys Arg Ile   |      |
|    | 55 60 65                                                          |      |
| 5  | ATT CCT TCA AGT ATG TTA ACT ATA TAT ACT CAA ATA CAA AAG GAT ATA   | 724  |
|    | Ile Pro Ser Ser Met Leu Thr Ile Tyr Thr Gln Ile Gln Lys Asp Ile   |      |
|    | 70 75 80                                                          |      |
| 10 | AAA AAC GGG AAT ATT GAC ACC GAA AAA TTA AGG AAA TAT GAG ATA GCA   | 772  |
|    | Lys Asn Gly Asn Ile Asp Thr Glu Lys Leu Arg Lys Tyr Glu Ile Ala   |      |
|    | 85 90 95 100                                                      |      |
| 15 | AAA GGA TTA ATG TCC GTT CCT TAT ATA TAT TTC TAATTTTTTC AATGATGTTA | 825  |
|    | Lys Gly Leu Met Ser Val Pro Tyr Ile Tyr Phe                       |      |
|    | 105 110                                                           |      |
|    | GTTGACTTCA AAAAGATGTG AAATCGATTA GCATTTTCAA AATTAGATTA AAAATACTAT | 885  |
| 20 | CTATATAAAA TAGAACTACT GATTTAAAGT ATTTATAAGA ATATAAAGTA GCAAATAACA | 945  |
|    | TGATAGACAC AATTAAGGAG CGACATTTTA TGGAAAATTT GAAATGGTAT TCGGGCGGGA | 1005 |
| 25 | ACGATAGAAA AAAAAAGCA ATG GCT ACT ATT ACT GAT TTG TTA AAC GAT TTA  | 1057 |
|    | Met Ala Thr Ile Thr Asp Leu Leu Asn Asp Leu                       |      |
|    | 1 5 10                                                            |      |
| 30 | AAA ATA GAC TTA GGT AAC GAA TCT CTA CAA AAT GTC TTA GAA AAT TAT   | 1105 |
|    | Lys Ile Asp Leu Gly Asn Glu Ser Leu Gln Asn Val Leu Glu Asn Tyr   |      |
|    | 15 20 25                                                          |      |
| 35 | CTT GAA GAA TTG GAA CAA GCA AAT GCT GCT GTT CCA ATT ATA TTA GGC   | 1153 |
|    | Leu Glu Glu Leu Glu Gln Ala Asn Ala Ala Val Pro Ile Ile Leu Gly   |      |
|    | 30 35 40                                                          |      |
| 40 | CGT ATG AAC ATA GAT ATC TCT ACA GCA ATC AGA AAA GAT GGT GTT ACT   | 1201 |
|    | Arg Met Asn Ile Asp Ile Ser Thr Ala Ile Arg Lys Asp Gly Val Thr   |      |
|    | 45 50 55                                                          |      |
| 45 | TTA TCA GAA ATT CAG TCT AAA AAA TTA AAA GAG CTG ATT TCA ATA TCC   | 1249 |
|    | Leu Ser Glu Ile Gln Ser Lys Lys Leu Lys Glu Leu Ile Ser Ile Ser   |      |
|    | 60 65 70 75                                                       |      |
| 50 | TAT ATT AAA TAT GGC TAT TAATTTAGTA TTAATAACAG TGTAGGATTG          | 1297 |
|    | Tyr Ile Lys Tyr Gly Tyr                                           |      |
|    | 80                                                                |      |
|    | ATTCAAATTA TTTGAATCAA AATTTATATA CAAATTTTAT TTATTTTGGG TCTTTAAATA | 1357 |
|    | ATTTTGTGTA AGTTCAAATT ATTTAAAGAT GAGTTAAAC TCTATCTTCG AAAAACATCA  | 1417 |
|    | CAAAATGTGA TGAAATTTGT CCCCAATTTT GGACCTTCAT GGTCCATTTT TTCGTTACAT | 1477 |
|    | CCATCGTCAC TAAACAAAGC ATTTTTAGTA AGGATTCATC AGATGGGAAT ACTACCTTAG | 1537 |

ATTTTGTGG CTTTCACAGC TGACAATGGA GGCCTTCAAT CACATTGGCG GTATAAATAA 1597  
TCCGGCGCAA ATCTGCTGAA TACTTGAAAA ATGTCGCTAA TTCTGCCCAG TTATCTATCA 1657  
5 ATTCTTCAAT TAATTTGTCT GCGTTTATAT CCATTTTCAT TCCCCTTTT TAATTTTCA 1717  
TTTTTTAGTT ACTTTAAACG GTTTAAAGCC TTAAGCACTT AGGCTTTAAT CTTTTTTCAC 1777  
TTGATCTAAT TATTTGAACT TCAGCATTTA TCTTTTGATT TATTCTTTTA GGGAATTGAC 1837  
10 CGAATAGGGA GATTTCTGT GAGTAGGCGC CAACGGTGGT GGCGGTCGGA GTCAGCCGAC 1897  
TCACAAGCTT 1907

15 (2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 66 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Asn Ser Val Lys Glu Leu Asn Val Lys Glu Met Lys Gln Leu His  
1 5 10 15  
20 Gly Gly Val Asn Tyr Gly Asn Gly Val Ser Cys Ser Lys Thr Lys Cys  
25  
30 Ser Val Asn Trp Gly Gln Ala Phe Gln Glu Arg Tyr Thr Ala Gly Ile  
35 35 40 45  
Asn Ser Phe Val Ser Gly Val Ala Ser Gly Ala Gly Ser Ile Gly Arg  
50 55 60  
40 Arg Pro  
65

(2) INFORMATION FOR SEQ ID NO:38:

45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 111 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Asp Ile Lys Ser Gln Thr Leu Tyr Leu Asn Leu Ser Glu Ala Tyr

1 5 10 15  
Lys Asp Pro Glu Val Lys Ala Asn Glu Phe Leu Ser Lys Leu Val Val  
20 25 30  
5 Gln Cys Ala Gly Lys Leu Thr Ala Ser Asn Ser Glu Asn Ser Tyr Ile  
35 40 45  
10 Glu Val Ile Ser Leu Leu Ser Arg Gly Ile Ser Ser Tyr Tyr Leu Ser  
50 55 60  
His Lys Arg Ile Ile Pro Ser Ser Met Leu Thr Ile Tyr Thr Gln Ile  
65 70 75 80  
15 Gln Lys Asp Ile Lys Asn Gly Asn Ile Asp Thr Glu Lys Leu Arg Lys  
85 90 95  
Tyr Glu Ile Ala Lys Gly Leu Met Ser Val Pro Tyr Ile Tyr Phe  
100 105 110

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 81 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Met Ala Thr Ile Thr Asp Leu Leu Asn Asp Leu Lys Ile Asp Leu Gly  
1 5 10 15  
35 Asn Glu Ser Leu Gln Asn Val Leu Glu Asn Tyr Leu Glu Glu Leu Glu  
20 25 30  
40 Gln Ala Asn Ala Ala Val Pro Ile Ile Leu Gly Arg Met Asn Ile Asp  
35 40 45  
Ile Ser Thr Ala Ile Arg Lys Asp Gly Val Thr Leu Ser Glu Ile Gln  
50 55 60  
45 Ser Lys Lys Leu Lys Glu Leu Ile Ser Ile Ser Tyr Ile Lys Tyr Gly  
65 70 75 80  
Tyr

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:  
(B) CLONE: JMc7

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCCAAGCTTC TGCTGTAAAT TATGGTAATG GTGTT

35

15

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 36 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: cDNA

25

(vii) IMMEDIATE SOURCE:  
(B) CLONE: KLR179

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCGCAAGCTT CTGCTCGGAC ACCAGAAATG CCTGTT

36

35 (2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: cDNA

45

(vii) IMMEDIATE SOURCE:  
(B) CLONE: KLR180

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GGCCAAGCTT GCCATTAAGT CTGGTTGCTA

30

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:  
(B) CLONE: MB32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AATTCGAGCT CGCCCAAATC

20

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 40 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:  
(B) CLONE: MB37

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TGAGTAATTT TCGGTGCAGC ACCTCCTACG ACTTGTTCTGA

40

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vii) IMMEDIATE SOURCE:  
(B) CLONE: RW58

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TACGCGCAAG AACAGACAAA ATC

23



(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 40 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:  
(B) CLONE: MB38

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TGAGTAATTT TCGGTGCAGC TCCTCCGTTA GCTTCTGAAA

40

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:  
(B) CLONE: MB39

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TACGAATTCTG AGCTCGCCC

19

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 76 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:  
(B) CLONE: MB42

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ATTTTCGGTG CAGCACCTCC AGAAACAGAA TCTAATTCAT TTAGAGTCAG AGTTCTCATA 60  
ATAACTTTCC TCTTTT 76

5 (2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 40 base pairs  
    (B) TYPE: nucleic acid  
10 (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- 15 (vii) IMMEDIATE SOURCE:  
    (B) CLONE: MB41

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TGAGTAATTT TCGGTGCAGC CATAATAACT TTCCTCTTTT 40

25 (2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 37 base pairs  
    (B) TYPE: nucleic acid  
30 (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- 35 (vii) IMMEDIATE SOURCE:  
    (B) CLONE: MB43

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATATCACGCC CTGAAGCACC TCCTACGACT TGTTCGA 37

(2) INFORMATION FOR SEQ ID NO:51:

- 45 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 36 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

(B) CLONE: MB44

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

5 AATTAAGCTT GGATCCTTCT GTGTGGATTG TCCAAT 36

(2) INFORMATION FOR SEQ ID NO:52:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

20 (vii) IMMEDIATE SOURCE:  
(B) CLONE: APO-1

25 (ix) FEATURE:  
(A) NAME/KEY: misc\_difference  
(B) LOCATION: replace(11..12, "")  
(D) OTHER INFORMATION: /standard\_name= "(any base)"

30 (ix) FEATURE:  
(A) NAME/KEY: misc\_difference  
(B) LOCATION: replace(16..17, "")  
(D) OTHER INFORMATION: /standard\_name= "(any base)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

35 AAAGATATTG GAAAGGATTG G 21